Online Appendix
Inventors among the "Impoverished Sophisticate"

Thor Berger* Erik Prawitz†

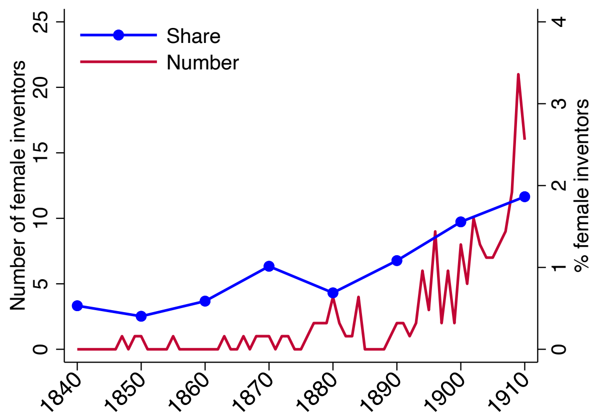
Contents

*Swedish Collegium for Advanced Study (SCAS), Uppsala University. Department of Economic History & Centre for Economic Demography, School of Economics and Management, Lund University. Centre for Economic Policy Research (CEPR). Research Institute of Industrial Economics (IFN), Stockholm. E-mail: thor.berger@ekh.lu.se

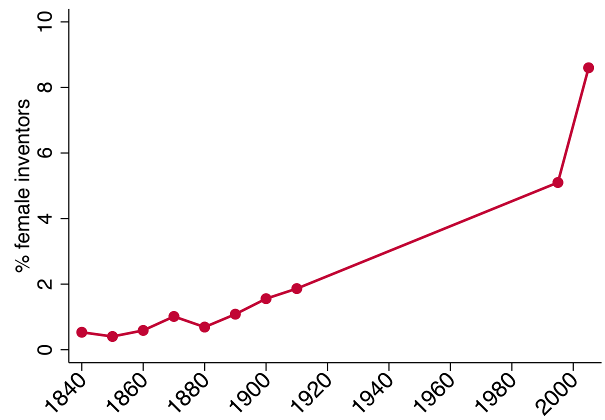
†Department of Economics and Statistics, Linnaeus University. Research Institute of Industrial Economics (IFN), Stockholm. E-mail: erik.prawitz@lnu.se

A Additional material

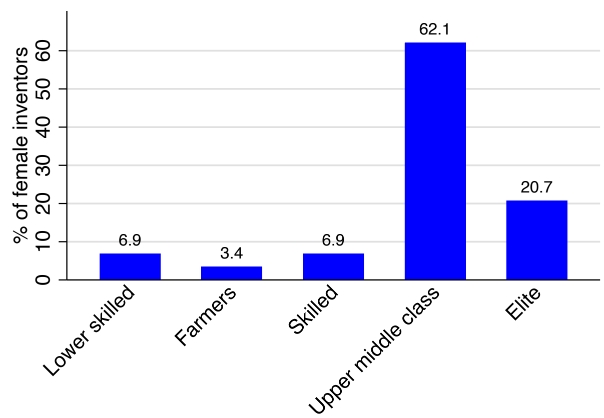
A.1 Demographic characteristics



(A) Female inventors, 1840–1914



(B) Female inventors, 1840–2005



(C) Social class of female inventors

Figure A.1:
Female inventors

Notes: A: The share of inventors that are identified as female between 1840 and 1914 among all active inventors and the number of female inventors that apply for a patent in each year. B: The same data as A paired with modern data from Jung and Ejermo (2014). C: The social class of female inventors using the *full inventor sample*. The different status categories are based on the HISCLASS social class scheme, as described in the main text.

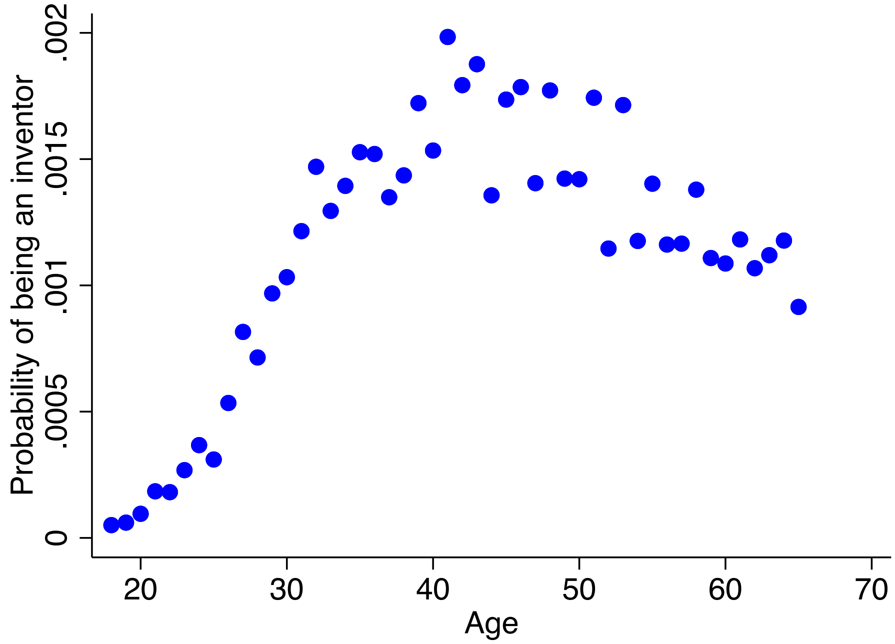


Figure A.2:
Probability to be an inventor by age.

Notes: The figure displays the probability to be an inventor for each age group (using the *census sample*).

A.2 Income, occupations, and social status

Table A.1: Most common occupational titles among inventors, 1885–1914.

	Freq.	Percent	Cum.
Elite			
Engineers	1283	46.72	46.72
General Manager	844	30.74	77.46
Officer	154	5.61	83.07
Mining Engineer	36	1.31	84.38
Building Architect	35	1.27	85.65
Upper Middle Class			
Working Proprietor	619	31.69	31.69
Production Supervisor or Foreman	297	15.21	46.90
Contractor	128	6.55	53.46
Bookkeeper	92	4.71	58.17
Production Manager	50	2.56	60.73
Skilled			
Machinery Fitter or Machine Assembler	260	18.39	18.39
Carpenter	220	15.56	33.95
Blacksmith	200	14.14	48.09
Tool and Die Maker	55	3.89	51.98
Watch and Clock Assembler or Repairer	55	3.89	55.87
Farmers			
General Farmer	224	93.72	93.72
Horticultural Farmer	14	5.86	99.58
Other Specialised Farmers	1	0.42	100.00
Lower Skilled			
Ship's Fireman	155	34.14	34.14
Metal Processor	51	11.23	45.37
Building Painter	45	9.91	55.29
Blacksmith	23	5.07	60.35
Dairy Product Processor	16	3.52	63.88
Unskilled			
Factory Worker	44	34.65	34.65
Worker	43	33.86	68.50
Labourer	8	6.30	74.80
Chimney Sweep	6	4.72	79.53
Farm Worker	6	4.72	84.25

Notes: The table shows the five most common occupational titles among Swedish inventors within each of the six broad social classes we examine (based on the HISCLASS social class scheme) using the *full inventor sample*.

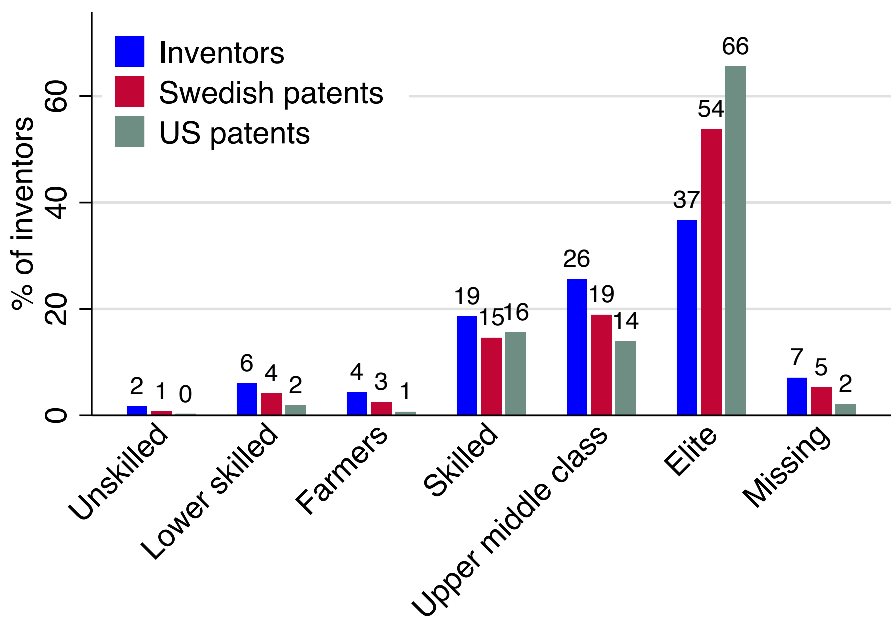


Figure A.3:
Social class of inventors including inventors with missing occupation

Notes: Distribution of social class among Swedish inventors (using the *full inventor sample*). The different status categories are based on the HISCLASS social class scheme, as described in the main text.

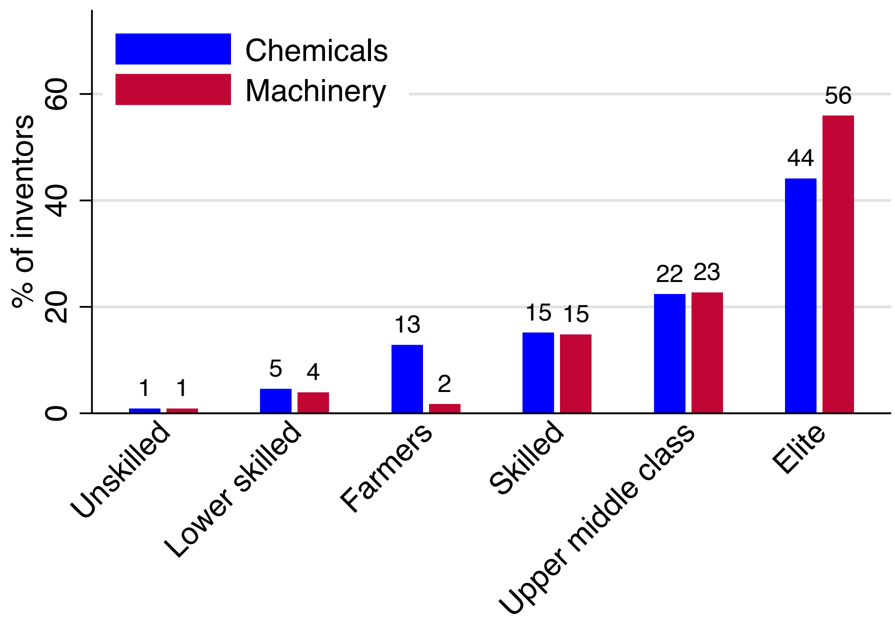


Figure A.4:
Social class of inventors and secrecy.

Notes: The figure displays the distribution of social class among Swedish inventors (using the *full inventor sample*) separately for machinery and chemical patents. The different status categories are based on the HISCLASS social class scheme, as described in the main text.

Secrecy. One potential explanation is that elite inventors selected into formal patenting due to other factors, e.g., better social networks or some form of institutional knowledge, whereas less skilled groups chose to invent in sectors protected by secrecy. To explore this possibility, we study to what extent elite inventors were similarly overrepresented in sectors better protected by secrecy. Under the assumption that inventive activity in these sectors would spill over to patenting, we can explore to what extent the skill background of inventors differs between these two sectors. Figure A.4 compares the social status distributions within chemical and machinery patents, as examples of sectors protected by high (chemical) and low (machinery) secrecy, showing a roughly similar pattern.[1] Although, we cannot rule out the importance of secrecy, this suggests that elite inventors were overrepresented also in sectors where formal patenting was less prominent.

[1]While secrecy became less prominent in the chemical industries towards the turn of the century (Moser, 2012), the bulk of patenting activity took place in a period marked by differential patenting rates (Moser, 2005).

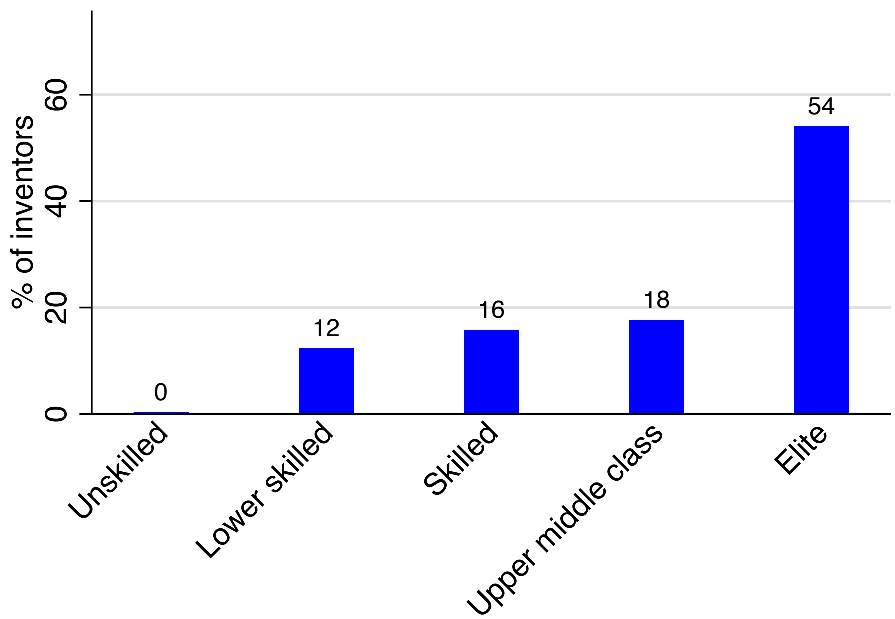


Figure A.5:
Social class of inventors not listed as patent holders.

Notes: The figure displays the distribution of social class among Swedish inventors not listed as patent holders (using the *full inventor sample*). The different status categories are based on the HISCLASS social class scheme, as described in the main text.

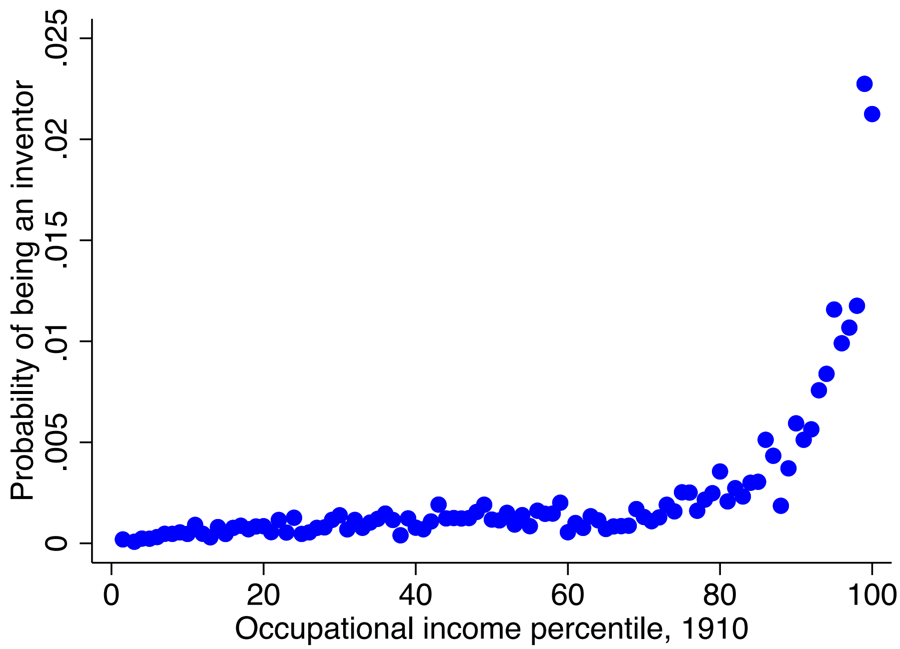


Figure A.6:
Income distribution of inventors

Notes: The figure displays a binned scatter plot of an indicator capturing whether an individual was an inventor and his occupational income score in 1910 (using the *census sample*). Observations are sorted into 100 groups of equal size and the circles indicate the mean probability of an individual being an inventor in each group.

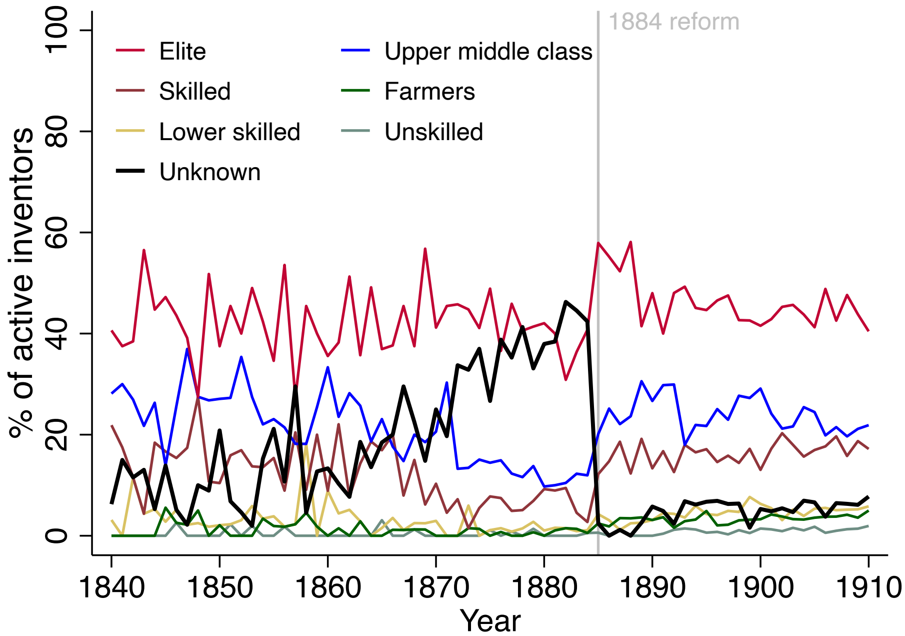
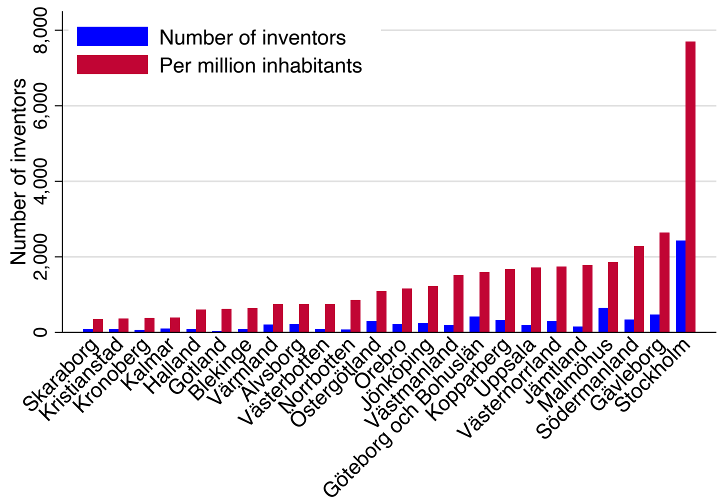


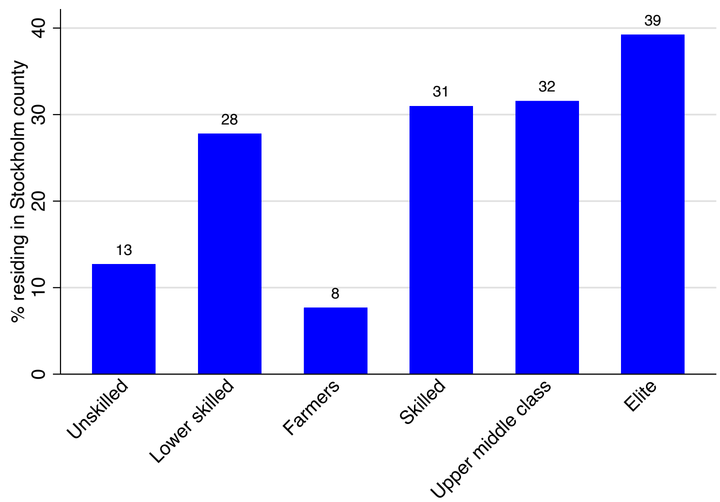
Figure A.7:
Social class of inventors during industrialization.

Notes: The figure displays the distribution of social class among active inventors granted at least one patent between 1840 and 1910. The different status categories are based on the HISCLASS social class scheme, as described in the main text. We include one observation for each inventor that patented at least once in each given year. We include and report the share of active inventors with missing occupational information.

A.3 The geography of inventors



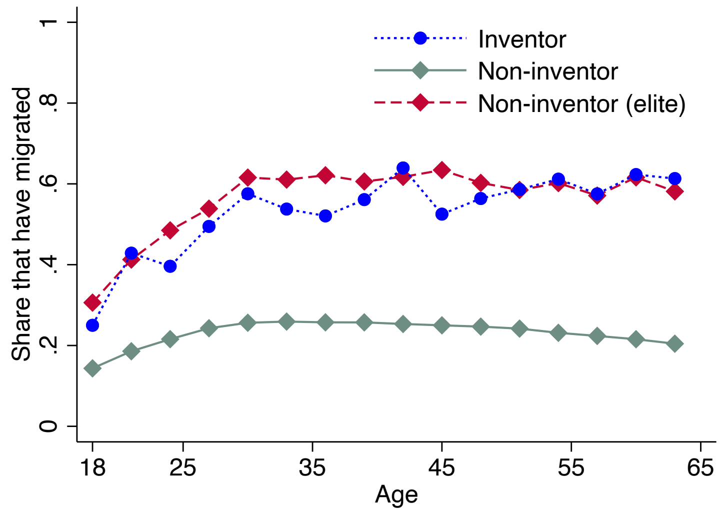
(A) Inventors by county



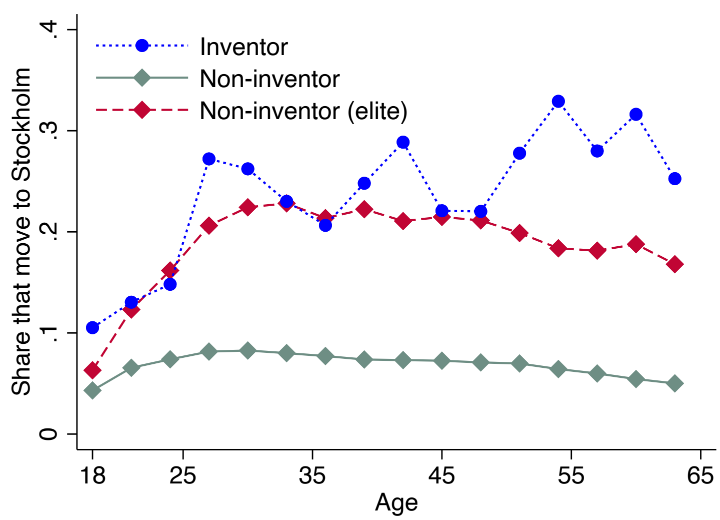
(B) Share of inventors residing in Stockholm

Figure A.8:
The geography of inventors

Notes: A: Number of inventors by county of residence denoted on the patent records. Inventors per million inhabitants is calculated based on population data from Statistics Sweden for 1880. B: The share of inventors that resided in Stockholm county by the social class of inventors. The different status categories are based on the HISCLASS social class scheme, as described in the main text. Both figures use the *full inventor sample.*



(a) Across counties

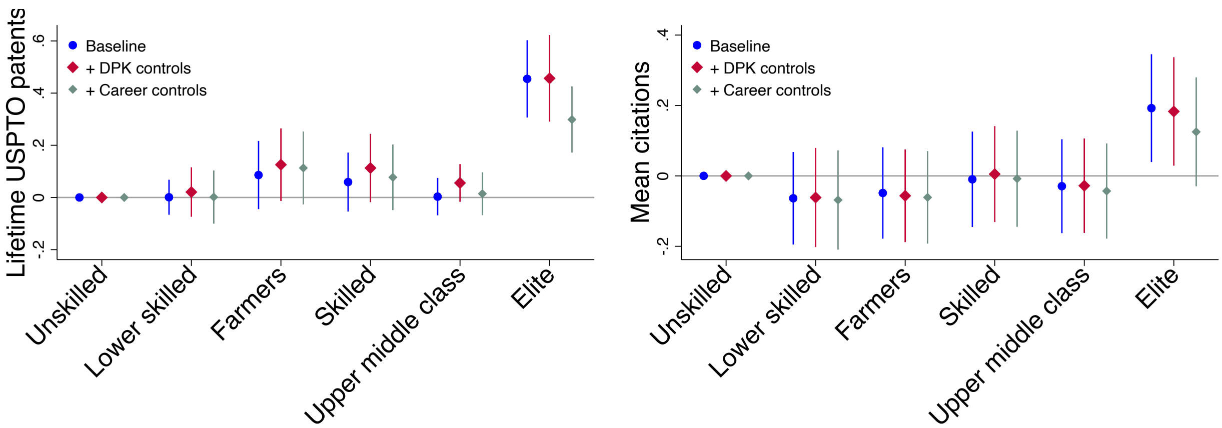


(b) To Stockholm

Figure A.9:
Geographic mobility of (non-)inventors over their life cycle

Notes: A: The share of (non-)inventors that reside in a different county than their county of birth in the 1910 census by their age in 1910. B: The share of (non-)inventors that reside in Stockholm county in 1910 by their age in 1910 (sample consists of those not born in Stockholm county). Both figures use the *census sample*.

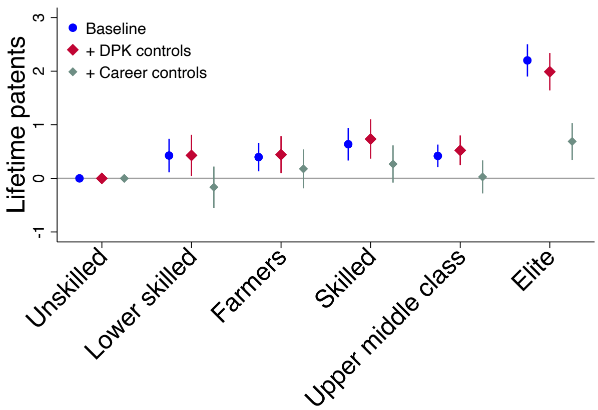
A.4 Inventor productivity



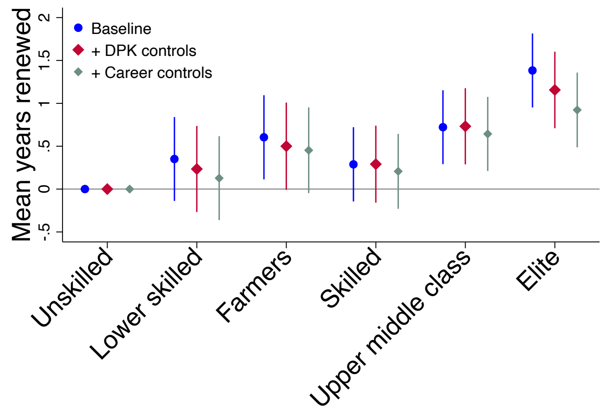
(A) Patent output (USPTO)   (B) Mean citations (USPTO)

Figure A.10:
Inventor output and quality by social class.

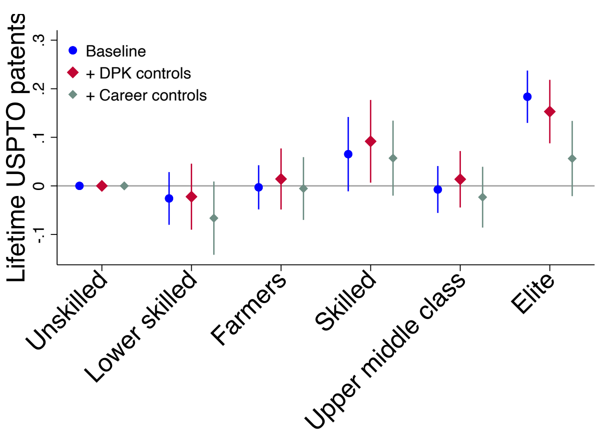
Notes: The figure displays point estimates and 95% confidence intervals from inventor-level OLS regressions of patent output and quality among inventors belonging to different social classes relative to inventors belonging to the unskilled class (using the *census sample*). A: The total number of granted USPTO patents over an inventor's lifetime. B: The average number of patent citations. The baseline regressions (denoted by blue circles) include controls for the first decade in which an inventor applied for a (subsequently granted) patent and the county of residence. Additional specifications add controls for the first (DPK) technology class an inventor patents in (red diamonds) and career length (teal diamonds).



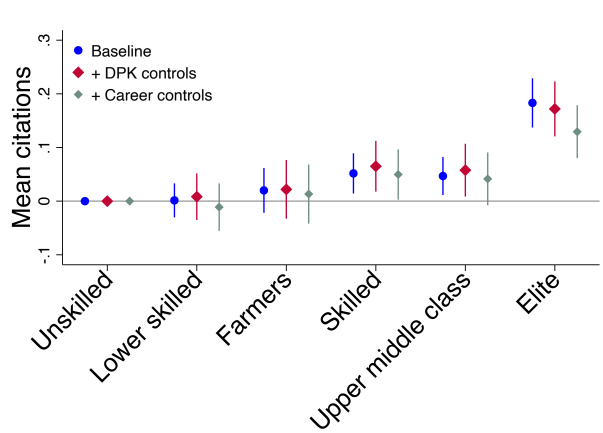
(A) Patent output (PRV)



(B) Patent quality (PRV



(C) Patent output (USPTO)



(D) Mean citations (USPTO)

Figure A.11:
Inventor output and quality by social class.

Notes: The figure displays point estimates and 95% confidence intervals from inventor-level OLS regressions of patent output and quality among inventors belonging to different social classes relative to inventors belonging to the unskilled class (using the *full inventor sample*). A: The total number of granted patents over an inventor's lifetime. B: The average number of years patent fees were paid per patent. C: The total number of granted USPTO patents over an inventor's lifetime. D: The average number of patent citations. The baseline regressions (denoted by blue circles) include controls for the first decade in which an inventor applied for a (subsequently granted) patent and the county of residence. Additional specifications add controls for the first (DPK) technology class an inventor patents in (red diamonds) and career length (teal diamonds).

A.5 Family background of inventors

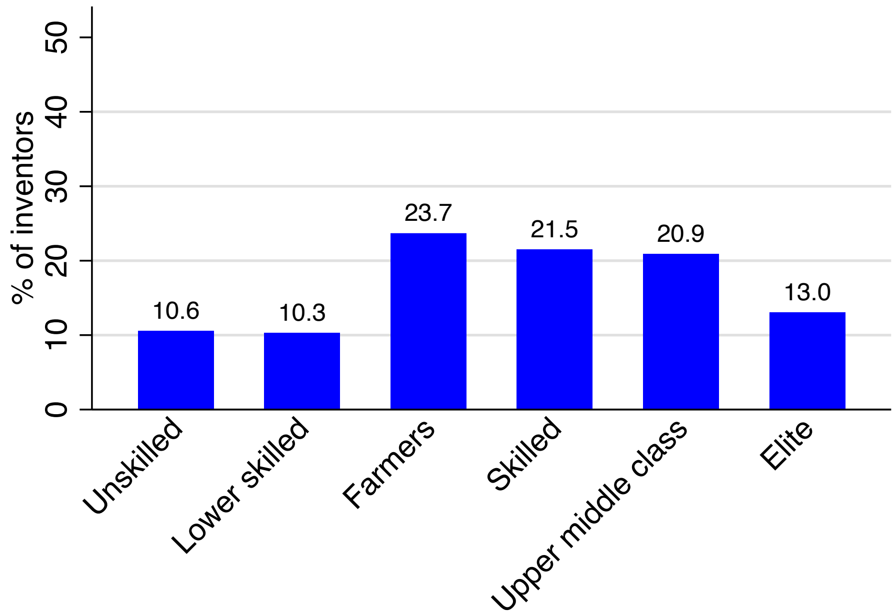
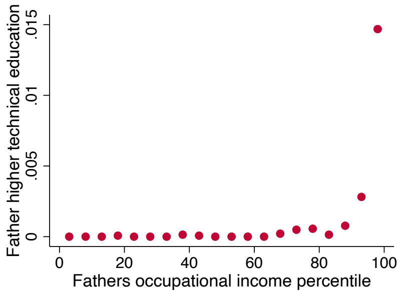
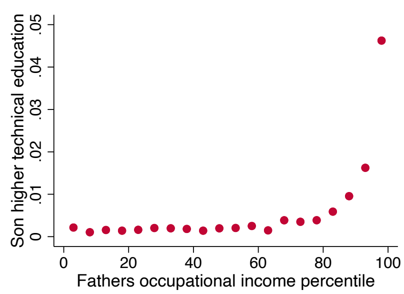


Figure A.12:
Social origins of inventors.

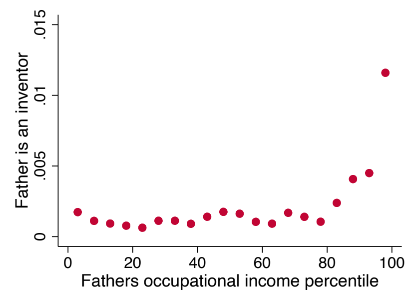
Notes: The figure displays the distribution of social class among fathers to inventors in our *linked father-son sample*. The different status categories are based on the HISCLASS social class scheme, as described in the main text.



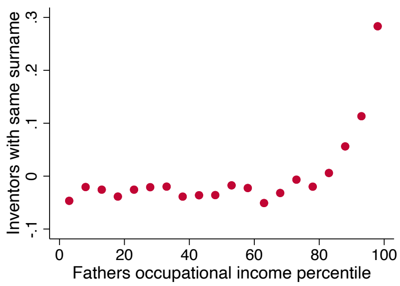
(A) Father: education



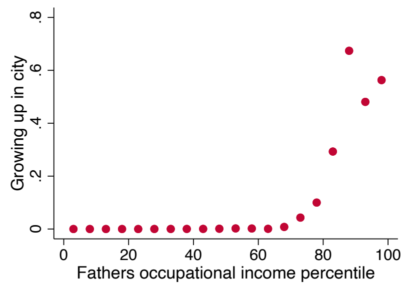
(B) Son: education



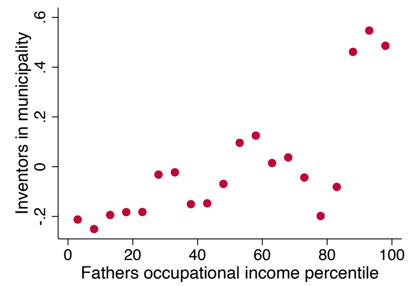
(C) Father: inventor



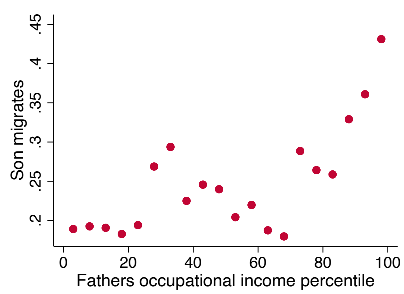
(D) Inv. with same surname



(E) Lives in urban area



(F) Inventors in municipality



(G) Son: migrates

Figure A.13:
Fathers income and outcomes for fathers and sons.

Notes: The figure displays binned scatter plots of outcomes for sons or fathers and the father's occupational income score in 1880 (using the *linked father-son sample*). Observations are sorted into 20 groups of equal size and the circles indicate the mean of the outcome within each group.

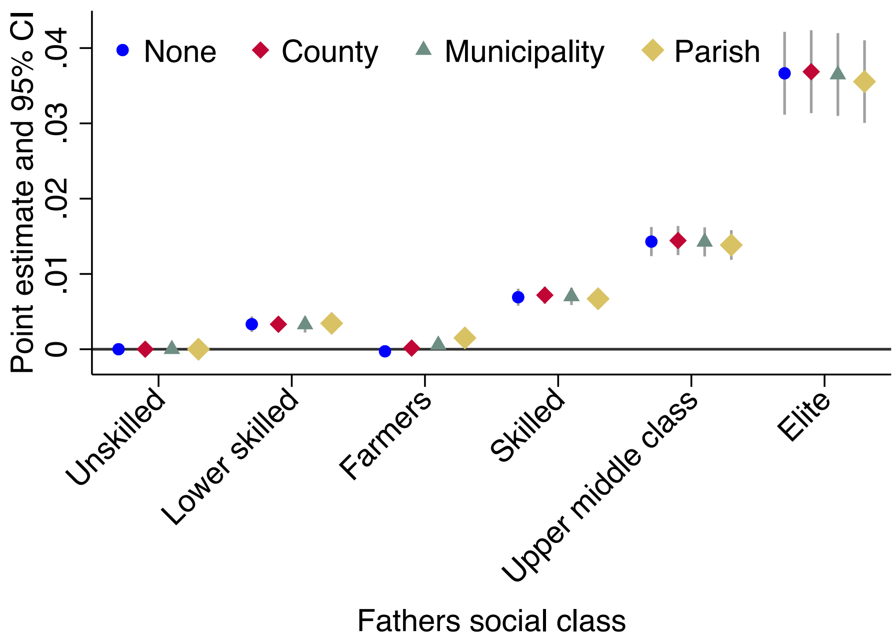


Figure A.14:
Family background and becoming an inventor - controlling for childhood location

Notes: The figure displays the relationship between father's social class and the probability of being an inventor in adulthood when including geographical fixed effects at different levels based on the *linked father-son sample.*

Table A.2: Who becomes an inventor - intensive margin

Dependent variable:	Star inventor (=1)					Number of patents					Years patents renewed				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Father's economic and social class															
Father top-10% (=1)	0.030***	0.024*	0.021	0.022	0.005	1.367***	1.283*	1.144	1.060*	0.481	0.239	0.193	0.171	0.100	0.015
	(0.011)	(0.014)	(0.013)	(0.015)	(0.013)	(0.418)	(0.695)	(0.719)	(0.587)	(0.491)	(0.256)	(0.259)	(0.262)	(0.300)	(0.319)
Father elite (=1)		0.020	-0.006	-0.006	-0.025		0.269	-0.586	-0.543	-1.199		0.148	0.006	0.026	-0.059
		(0.025)	(0.025)	(0.025)	(0.025)		(1.323)	(1.165)	(1.135)	(1.187)		(0.241)	(0.242)	(0.242)	(0.252)
Family exposure to innovation															
Father higher technical education (=1)			0.098	0.096	0.054			2.895	2.802	1.210			0.416	0.385	-0.008
			(0.068)	(0.067)	(0.075)			(2.436)	(2.493)	(3.016)			(0.611)	(0.625)	(0.649)
Father inventor (=1)			0.029	0.028	0.028			2.937	2.858	2.796			0.488	0.453	0.377
			(0.039)	(0.040)	(0.043)			(2.176)	(2.242)	(2.376)			(0.470)	(0.484)	(0.485)
Inventors with same surname, pre-1880			0.008***	0.008***	0.008**			0.221**	0.219**	0.197			0.041	0.041	0.038
			(0.003)	(0.003)	(0.003)			(0.103)	(0.104)	(0.121)			(0.035)	(0.035)	(0.031)
Local exposure to innovation															
Lives in urban area, 1880 (=1)				-0.006	-0.006				-0.047	-0.064				0.117	0.130
				(0.016)	(0.016)				(0.551)	(0.544)				(0.309)	(0.305)
Inventors in municipality, pre-1880				0.006*	0.004				0.242***	0.168*				0.040	0.018
				(0.003)	(0.004)				(0.085)	(0.092)				(0.060)	(0.057)
Son's education and location															
Son higher technical education (=1)					0.117***					4.313***					0.938***
					(0.028)					(1.456)					(0.288)
Migrant, 1880-1910 (=1)					0.012					0.225					-0.270
					(0.010)					(0.537)					(0.292)
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Childhood county FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1456	1456	1456	1456	1456	1456	1456	1456	1456	1456	1448	1448	1448	1448	1448
Mean dep. var.	0.041	0.041	0.041	0.041	0.041	3.001	3.001	3.001	3.001	3.001	4.924	4.924	4.924	4.924	4.924

Notes: The table reports individual-level OLS regressions using the *linked father-son sample* between the 1880 and 1910 census. Individual controls include cubic functions in the age of the father in 1880 and the son in 1910, respectively. All regressions include a full set of fixed effects for county of residence in 1880. Standard errors are given in parentheses and are clustered at the county level. *** - $p < 0.01$, ** - $p < 0.05$, * - $p < 0.1$.

Table A.3: Intergenerational income mobility among (non-)inventors: IGE and rank-rank estimates.

Dependent variable:	Son's income (ln)			Son's income rank		
	Inventors	Non-inventors	All	Inventors	Non-inventors	All
	(1)	(2)	(3)	(4)	(5)	(6)
Father's income (ln)	0.511***	0.569***	0.569***			
	(0.024)	(0.003)	(0.003)			
Inventor (=1)			0.655***			21.133***
			(0.164)			(2.069)
Income (ln) × Inventor (=1)			-0.058**			
			(0.024)			
Father's income rank				0.388***	0.476***	0.476***
				(0.026)	(0.002)	(0.002)
Rank × Inventor (=1)						-0.084***
						(0.025)
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1456	283188	284644	1456	283188	284644
Mean dep. var.	7.06	6.65	6.65	74.09	50.35	50.47

Notes: The table reports individual-level OLS regressions using the *linked father-son sample* between the 1880 and 1910 census. Individual controls include cubic functions in the age of the father in 1880 and the son in 1910, respectively. Standard errors are given in parentheses and are clustered at the father level. *** - $p < 0.01$, ** - $p < 0.05$, * - $p < 0.1$.

Table A.4: Invention and intergenerational income mobility

Dependent variable:	Son's ln occ. income score, 1910			
	(1)	(2)	(3)	(4)
Inventor (=1)	0.426***	0.103***		
	(0.019)	(0.018)		
Inventor: pre-1910 (=1)			0.151***	
			(0.026)	
Inventor: post-1910 (=1)			0.053**	
			(0.024)	
Inventor: 1 patent (=1)				0.082***
				(0.023)
Inventor: 2-9 patents (=1)				0.116***
				(0.029)
Inventor: 10+ patents (=1)				0.260***
				(0.087)
Individual controls	Yes	Yes	Yes	Yes
Father FE	No	Yes	Yes	Yes
Observations	140448	140448	140448	140448
Mean dep. var.	6.65	6.65	6.65	6.65

Notes: The table reports individual-level OLS regressions using the *linked father-son sample* between the 1880 and 1910 census. The dependent variable is a son's ln occupational income score in 1910. Individual controls correspond to a cubic in sons' age in 1910. We restrict all samples to sons where we observe at least one brother. Standard errors are given in parentheses and are clustered at the father level. *** - $p < 0.01$, ** - $p < 0.05$, * - $p < 0.1$.

Table A.5: Invention and intergenerational income mobility

Dependent variable:	Upward mobility (=1)			
	(1)	(2)	(3)	(4)
Inventor (=1)	0.109***	0.081***		
	(0.019)	(0.019)		
Inventor: pre-1910 (=1)			0.106***	
			(0.026)	
Inventor: post-1910 (=1)			0.055**	
			(0.027)	
Inventor: 1 patent (=1)				0.065**
				(0.025)
Inventor: 2-9 patents (=1)				0.100***
				(0.029)
Inventor: 10+ patents (=1)				0.136
				(0.085)
Individual controls	Yes	Yes	Yes	Yes
Father FE	No	Yes	Yes	Yes
Observations	140448	140448	140448	140448
Mean dep. var.	0.36	0.36	0.36	0.36

Notes: The table reports individual-level OLS regressions using the *linked father-son sample* between the 1880 and 1910 census. The dependent variable is an indicator taking the value one if a son's occupational income rank in 1910 surpasses that of his father in 1880. Individual controls correspond to a cubic in sons' age in 1910. We restrict all samples to sons where we observe at least one brother. Standard errors are given in parentheses and are clustered at the father level. *** - $p < 0.01$, ** - $p < 0.05$, * - $p < 0.1$.

A.6 Occupational mobility

We here examine whether inventors also exhibit more occupational mobility asking whether inventors were more likely to transition out of their fathers social class. Here we focus on the same six broad social groups used throughout the paper: elite, upper middle class, skilled, farmers, lower skilled, and unskilled. Inventors were much more likely to transition out of their fathers social class: 67 percent of inventors are observed in a different occupational group than their fathers, which can be compared to 58 percent among non-inventors.[2]

To paint a richer picture of the occupational origins and destinations of inventors, Appendix Table A.6 displays the occupational origins and destinations for sons based on their fathers occupation. Here we observe children's occupational groups in the 1910 census, while their fathers occupations are observed in 1880. We observe more persistence among inventors in the top of the distribution: 63 percent of inventor sons born to elite fathers remain in elite occupations as adults, which can be compared to 39 percent among non-inventors. At the same time, we observe more mobility in the bottom of the distribution. For example, only 16 percent of inventor sons born to fathers with an unskilled occupation remain in an unskilled occupation in adulthood.

To more formally measure rates of relative occupational mobility, we also estimate Altham statistics as common in the historical literature (Long and Ferrie, 2013; Pérez, 2019; Berger et al., 2023). The Altham statistic summarizes all the odds ratios in a mobility table, which reflect the relative chances of reaching a given occupational standing for sons from different origins. We then compare the mobility table of inventors and non-inventors ($\mathbf{P}$) to a table ($\mathbf{I}$) where the occupational attainment of sons is independent of their fathers. The Altham $d(\mathbf{P}, \mathbf{I})$ statistic ranges between 0 and infinity, where a larger statistic corresponds to a greater departure from the case of full mobility (i.e., less mobility). The Altham $d(\mathbf{P}, \mathbf{I})$ statistic is 58.5 ($p = 0.000$) and 49.8 ($p = 0.000$) for non-inventors and inventors respectively, which indicates a higher degree of intergenerational occupational mobility among inventors.

[2] The relatively higher mobility rates among inventors occurred against a backdrop of high rates of both absolute and relative occupational mobility among the Swedish population before World War I. Berger et al. (2023) shows that late-19th century Sweden exhibits higher intergenerational occupational mobility than other European countries and that mobility rates are closer to those observed in the highly mobile Americas (Long and Ferrie, 2013; Pérez, 2019).

Table A.6: Occupational father-son transitions for (non-)inventors, 1880–1910

Panel A. Inventors

	Son's occupation						
Father's occupation	Elite %	Upper middle class %	Skilled %	Farmers %	Lower skilled %	Unskilled %	Total %
Elite	63	23	10	2	2	0	100
Upper middle class	39	39	12	3	5	2	100
Skilled	19	28	35	1	9	8	100
Farmers	13	27	20	21	12	6	100
Lower skilled	9	27	24	6	23	11	100
Unskilled	8	21	29	9	17	16	100
Total	25	29	22	8	10	6	100
N	369	418	317	111	149	92	1,456

Panel B. Non-inventors

	Son's occupation						
Father's occupation	Elite %	Upper middle class %	Skilled %	Farmers %	Lower skilled %	Unskilled %	Total %
Elite	39	37	10	7	4	4	100
Upper middle class	11	41	13	10	12	13	100
Skilled	3	14	39	8	18	18	100
Farmers	1	8	12	49	13	18	100
Lower skilled	1	12	19	13	34	22	100
Unskilled	1	10	19	12	26	33	100
Total	2	12	17	30	18	21	100
N	6,737	33,990	48,157	85,369	50,524	58,411	283,188

Notes: The table displays occupational transitions for sons relative to their fathers using the *linked father-son sample*. Each row corresponds to the occupational group of fathers observed in the 1880 census. Each column corresponds to the occupation of sons observed in the 1910 census.

B Data appendix

B.1 Patent data

The patent data draws on a large database covering the whole population of granted Swedish patents 1746–1945.[3] It has been compiled using the following sources:

- *Kommerskollegium, Ingående diarier över patent, 1820-1884* (Swedish National Archive)
- *Bidrag till Sveriges officiella statistik (BiSOS) D: Fabriker och manufakturer, 1860-1884* (Statistics Sweden)
- *Förteckning över patenter beviljade i Sverige och Norge 1866–1875* (L. A. Groth & Co Patent Agency, Stockholm: 1876)
- *Patent- och registreringsverkets registratur, 1885-1914* (Swedish Intellectual Property Office)

The registers were stored in large hand-written ledgers. To minimize data entry errors and for more effective and systematic storage, a relational database structure was created and data entry performed through a structured and standardized template using a database software. Each patentee and inventor were given a unique identifier. To identify individuals across multiple patents, trained research assistants have created hand links using the full information in the hand-written ledgers, including name, occupation, address, patent agents, co-patentees and co-inventors, as well as patent type.

B.2 Linking inventors to the census

We here describe how we link the inventors from our patent data (A) to the 1910 census (B). To link inventors across time we compare first and last names for individuals of the same sex in the two data sets. Since spelling variations occur in the two data sets and since the inventor data sometimes do not provide a full list of first names, we link individuals using string distance metrics common in the literature. In particular, we make use of the Jaro-Winkler string distance metric, which measures name similarity on a scale between 0 (no similarity) and 1 (full similarity).

Moreover, while we do not know the birthyear of our inventors, we assume that inventors are at least 15 years of age when filing for a patent. Below we describe the used record linkage algorithm in detail.

[3] The database has been produced by a group of researchers at the Department of Business Administration at Uppsala University in collaboration with the Patent and Registration Office (PRV) during the period 2017–2021. See https://svenskahistoriskapatent.se for additional information.

1. We consider two records, an inventor (X) and a census individual (Y) as a match if they are the only pair with the same sex and the exact same names among candidates with a birth year in the census that is at least 15 years before the patent application.

2. If there is no single candidate above, we proceed by comparing names that has undergone a limited cleaning in terms nobiliary particles, suffixes, and a few common Swedish language spelling variations.[4]

3. If there is no single candidate above, we proceed by establishing links using name similarity. We consider a pair of individuals of the same sex as a match if the last names have a similarity of at least 0.9, the mean similarity of first names is at least 0.9, and there is no closely competing candidate. For the latter, we impose that the pair has the highest mean of the last name and first name JW scores, and it is at least 0.1 JW score units higher than the candidate with the second highest mean. To compute the similarity between first names without imposing any order of first names, we calculate the mean for the n number of first name pairs with the highest JW score, where n is equal to the number of first names in the record with the least number of first names in the pair.[5]

4. As a last step, we perform step 3 again after discarding candidates that are residing in a different county in 1910 than in their modal patent application (or first application if a modal is not applicable).

In total, we link 32.3 percent of inventors to the 1910 census. After discarding a few duplicates in terms of census individuals, we find ourselves with 3,236 inventors. Roughly 57 percent are established in step 1 above, 14 percent in step 2, 4 percent in step 3, and 25 percent in step 4. For these established links, the mean JW score is 0.97 for first names and 0.99 for last names.

[4]We make the following corrections: (i) V for W, (ii) K for C if C is followed by the vocals A or O, by T, or if C is the terminal character, (iii) V for F if preceded by A or O, (iv) S for double SS, (v) L for HL, and (vi) K for Q if followed by V.

[5]To exemplify, a census individual with the first names CARL GUSTAF PATRIK and an inventor with the single first name GUSTAF is given a JW score of 1, since the inventor has only one (1) first name (i.e. n=1), and since GUSTAF—GUSTAF is the pair with the highest JW score.

References

Berger, Thor, Per Engzell, Björn Eriksson, and Jakob Molinder, "Social mobility in Sweden before the welfare state," *The Journal of Economic History*, 2023, *83* (2), 431–463.

Jung, Taehyun and Olof Ejermo, "Demographic patterns and trends in patenting: Gender, age, and education of inventors," *Technological Forecasting and Social Change*, 2014, *86*, 110–124.

Long, Jason and Joseph Ferrie, "Intergenerational occupational mobility in Great Britain and the United States since 1850," *American Economic Review*, 2013, *103* (4), 1109–37.

Moser, Petra, "How do patent laws influence innovation? Evidence from nineteenth-century world's fairs," *American Economic Review*, 2005, *95* (4), 1214–1236.

_ , "Innovation without Patents: Evidence from World's Fairs," *The Journal of Law & Economics*, 2012, *55* (1), 43–74.

Pérez, Santiago, "Intergenerational occupational mobility across three continents," *The Journal of Economic History*, 2019, *79* (2), 383–416.